

УДК 004.06

ПРИМЕНЕНИЕ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ ДЛЯ АНАЛИЗА И СИНТЕЗА ИНФОРМАЦИОННЫХ ПРОЦЕССОВ

Письмеров А.М (Университет ИТМО)

Научный руководитель – к.т.н., доцент Муромцев Д.И.
(Университет ИТМО)

В работе рассматривается приложение методов векторных представлений к задаче генерации информационных процессов на основе их журнального описания и анализируется актуальность их использования при анализе информационных процессов. Предварительная кластеризация журналов событий может помочь выяснить, каким процессам принадлежат определенные события в журналах. Поскольку кластеризация логов — не всегда простая задача, предварительный переход от логов к векторным представлениям событий может стать важным шагом в решении задач интеллектуального анализа процессов. В этой работе я применяю различные методы векторных представлений к набору данных журналов событий. Переходя к векторам и применяя методы кластеризации, я повышаю эффективность интеллектуального анализа процессов.

Введение.

Развитие информационных технологий приводит к увеличению количества информационных процессов. Большинство таких процессов описывается в различных журналах путем логгирования каждого этапа этого процесса. Задача генерации представлений и моделей реальных процессов на основе их журнального описания возникает все чаще. Однако существующие подходы не обеспечивают достаточной степени точности генерируемых моделей, особенно при сложных структурах описываемого процесса.

Основная часть.

Предложен подход для решения задачи генерации информационных процессов на основе их цифрового следа. В данном подходе предлагается использовать векторные представления для отдельных событий и связанных цепочек этих событий. Применение метода векторного представления позволит представить задачу поиска отдельных процессов в как более привычную для области машинного обучения, а именно как задачу кластеризации векторов.

Таким образом, задача сводится к стандартному поиску кластеров. В рамках данной задачи каждый сформированный кластер представляет собой отдельный самостоятельный процесс. Такое представление возможно за счет использования метрик оценивающих схожесть векторных представлений различных цепочек событий. Более того, векторные представления позволяют использовать уже известные метрики для оценки качества кластеризации.

Сравнение нескольких методов векторных представлений произведенных в рамках данной работы позволяет считать развитие данного подхода перспективным. Применение векторных представлений расширяет возможности исследования различных методов для решения поставленной задачи, особенно на архитектурно более сложных процессах.

Выводы. В рамках работы были исследованы методы векторного представления, применимые для анализа, синтеза информационных процессов на основе их цифрового следа. Смоделирован и проведен эксперимент. С помощью эксперимента продемонстрировано, что применение методов векторных представлений могут быть использованы для решения задачи генерации модели информационного процесса. При этом, предложенные подходы демонстрируют достаточно высокий уровень точности построенных моделей.